The nilpotent (n,n(n+1)/2) sub-Riemannian problem

Felipe Monroy-Pérez, Alfonso Anzaldo-Meneses

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

The nilpotent $(n, n(n+1)/2)$ *sub-Riemannian problem*

Felipe Monroy-Pérez — Alfonso Anzaldo-Meneses

N° 4990

Novembre, 2003

THÈME 4

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--4990--FR+ENG

The nilpotent $(n, n(n+1)/2)$ sub-Riemannian problem

Felipe Monroy-Pérez * , Alfonso Anzaldo-Meneses †

Thème 4 — Simulation et optimisation
de systèmes complexes
Projet MIAOU

Rapport de recherche n° 4990 — Novembre, 2003 — 20 pages

Abstract: In this paper we study the sub-Riemannian geodesic problem determined by a left invariant distribution of dimension n in a nilpotent Lie group G of dimension $n(n+1)/2$. The problem is formulated as an left invariant optimal control sytem. We describe a group action that leaves invariant the system. We apply Pontryagin maximum to derive necessary conditions for the optimality of the geodesics. For n odd abnormal extremals do exists, but they turn out to be non-strictly abnormal. The adjoint equation for the normal extremals is explicitly integrated. At the end, we specialize our results in some of the known low dimensional cases, namely, the Heisenberg $(2,3)$ case and the $(3,6)$ case.

Key-words: Nilpotent Lie algebras, optimal control, extremal curves.

* On sabbatical leave from the Basic Sciences Department of the Universidad Autonoma Metropolitana Azcapotzalco. Financially supported by the CONACYT, México.

† Basic Sciences Department of the Universidad Autonoma Metropolitana Azcapotzalco.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : +33 4 92 38 77 77 — Télécopie : +33 4 92 38 77 65

Le problème nilpotent sous-Riemannien $(n, n(n+1)/2)$

Résumé : Dans cet article, nous étudions le problème des géodésiques déterminées par une distribution invariante à gauche de dimension n sur un groupe de Lie nilpotent de dimension $n(n+1)/2$. Nous formulons le problème comme un problème de commande optimale invariant à gauche. Nous décrivons l'action d'un groupe sous laquelle le système est invariant. En appliquant le principe du maximum de Pontryaguine nous obtenons des conditions nécessaires pour l'optimalité des géodésiques. Pour n pair, il y a des extrémales anormales, mais elles ne sont pas strictement anormales. Les équations adjointes pour les extrémales normales sont integrées explicitement. Finalement, nous appliquons spécifiquement les résultats aux cas $(2,3)$ et $(3,6)$

Mots-clés : algèbres de Lie nilpotentes, commande optimal, courbes extrémales.

1 Introduction

This paper deals with left invariant distributions of dimension n defined on nilpotent Lie groups of dimension $n(n+1)/2$. Nilpotent Lie groups and algebras are important in the theory of distributions and in geometric nonlinear control theory. For control systems defined by families of vector fields, one expects to find equivalent families that generate nilpotent Lie algebras, exponential coordinates then can be used for analyzing the induced geometry at the level of the base manifold. Notions such as nilpotent approximations, nilpotentizations, and Carnot groups point out in that direction.

The study of length minimizing trajectories for left invariant distributions defined on nilpotent Lie groups has been in the literature in the last decades. Its treatment goes from the theory of hypoelliptic operators in mathematical physics to the sub-Riemannian geometry and geometric nonlinear control theory settings. Back in the mid-seventies, within the framework of sub-elliptic estimates for fundamental solutions of hypoelliptic operators, B. Gaveau establishes in [1], the relationship between what he calls horizontal variational problem and the problem of deterministic optimal control problem. In particular, he studies the system of bicharacteristics for a natural hypoelliptic operator in a free nilpotent Lie group of rank 2.

A thorough refinement of Gaveau's work was given later on by R. Brockett in [2], where he establishes the seminal ideas of sub-Riemannian geometry and tackles the issue of abnormal extremals. Although the existence of such extremals was finally settled down by R. Montgomery in [3]. There is an abundant literature around abnormal minimizers, and its significance in both, nonlinear control and Sub-Riemannian geodesic problem, see for instance [4], and [5]. We refer the reader to the recently published book [6] for a nice recount.

In the aforementioned reference, R. Brockett studies the geodesic problem associated to the optimal control problem given by the differential system $(\dot{x}, \dot{Y}) = (u, xu^T - ux^T)$ with $(x, Y) \in \mathbb{R}^n \times \mathrm{SO}_n$, boundary conditions $x(0) = x(1) = 0, Y(0) = 0, Y(1) = Y$, and cost functional

$$\eta = \int_0^1 \langle \dot{x}, \dot{x} \rangle \, dt.$$

Using the Lagrange multipliers technique he argues that optimal controls satisfy $\dot{u} + \Lambda u = 0$, for certain skew-symmetric matrix Λ, then he provides an explicit expression for the associated distance in terms of the eigenvalues of the matrix Λ. Brockett's argumentation has been further clarified by W. Liu and H. Sussmann in [5], whom elucidate the implications of this result regarding the existence of strictly abnormal extremals that cannot be minimizers.

The analysis of nonlinear control systems for which the control parameters appear linearly, is greatly simplified if the system is transformed into a system given by vector fields generating a nilpotent Lie algebra. The problem of nilpotent approximation for control systems of the form $\dot{x} = X_0(x) + u_1 X_1(x) + \cdots + u_k X_k(x)$, consists in finding conditions under which the system is equivalent to $\dot{x} = Y_0(x) + u_1 Y_1(x) + \cdots + u_k Y_k(x)$ and the Lie algebra generated by $\{Y_0, Y_1, \ldots, Y_k\}$ is nilpotent, see [7].

Nilpotent approximations have been also studied within the setting of sub-Riemannian geometry. For instance, the tangent space of a sub-Riemannian manifold (M, Δ) is defined by means of the nilpotent approximation of the distribution Δ. If $\Delta = \{X_1, \ldots, X_k\} \subset TM$ is the bracket generating distribution defining the sub-Riemannian structure, then using the canonical flag $\Delta \subset \Delta^2 \subset \cdots \subset \Delta^r = TM$, with $\Delta^j = [\Delta, \Delta^{j-1}]$, one assigns to each X_i an order ≥ -1. Then each X_i is expanded in series of homogeneous vector fields $X_i = X_i^{(-1)} + X_i^{(0)} + X_i^{(1)} + X_i^{(2)} + \cdots$, where $X_i^{(s)}$ has degree s. Defining $\widehat{X}_i = X_i^{(-1)}$ one gets the nilpotent approximation $\widehat{\Delta} = \{\widehat{X}_1, \ldots, \widehat{X}_k\}$. The distribution $\widehat{\Delta}$ generates a nilpotent Lie algebra, it is bracket generating and the sub-Riemannian distance is finite for each pair of points in M. Furthermore, in (privileged) coordinates, the system $\dot{x} = \sum_{i=1}^{k} u_i \widehat{X}_i(x)$ takes the form $\dot{z}_j = \sum_{i=1}^{k} u_i\, f_{ij}\ j = 1, \ldots, \dim M$, where the functions f_{ij} are homogeneous polynomials, for details see [8].

This nilpotentization is the closest sub-Riemannian geometric object that corresponds to the Riemannian tangent space. That is, generally speaking, sub-Riemannian manifolds are not infinitesimally Euclidean but rather modeled by simply connected Lie groups with finite dimensional nilpotent and graded Lie algebras, the so-called Carnot groups, see [9] .

The nilpotent approximation is deep-rooted on the general notion of equivalence of differential systems, which goes back to Élie Cartan [10]. The method of equivalence and the framework of exterior differential systems [11], have been brought into the control theory literature by R. Gardner [12] and co-workers. Such techniques have been successfully applied in the study of certain non-holonomic control systems encoded in kernels of Pfaffian systems, see for instance [13].

In this paper we study an optimal control problem on a nilpotent Lie group for which the dynamics is defined by means of distribution satisfying some general Lie commuting relations. We study this problem through the application of the Pontryagin Maximum Principle and the associated Hamiltonian formalism, see [14]. In particular we shall be interested in the integrability of the resulting Hamiltonian system.

Integrability theory of Hamiltonian systems associated to nonholonomic distributions, has a long history that goes back to the foundations of nonholonomic mechanics and thermodynamics. In contrast with the classical case, these systems appear usually as constrained Hamiltonian systems where the constraints play a central role in the geometry of the solutions. We refer the reader to the survey by A.M. Vershik and V.Ya. Gershkovich, [15], where they present an authorized recollection of nonholonomic dynamical system and the geometry of distributions along with a concise historical recount.

The integrability of the three dimensional contact case and its higher dimensional generalization is well known, [2], [16], whereas integrability of the non contact cases appear somehow interspersed in the literature and is far to be complete. The three-dimensional non contact case, the so-called Martinet case, has been extensively studied by B. Bonnard and co-workers, see [17], whereas the semi-contact case has been treated by A. Agrachev and co-workers [18], and by G. Charlot [19].

The four dimensional case, the Engel case, has been studied by A.M. Vershik and O. Granichina [20] and by Y.L. Sachkov in [21]. The so-called Cartan five dimensional case, originally studied by E. Cartan himself in [22], is discussed by R. Brockett and L. Dai in [23] within the framework of control theory and nonholonomic mechanics (rolling without slipping or twisting). In connection with the study of rigidity of curves and abnormality of extremals in the calculus of variations, R. Bryant and L. Hsu study in [24] an equivalent system in $\mathbb{R}^5$, using the formalism of exterior differential systems. In the aforementioned reference, Y.L. Sachkov also explicitly calculates the symmetries of this case. Some integrability aspects of the nilpotent $(3,6)$ sub-Riemannian problem are also discussed by O. Myasnichenko in [25], in the context of the geometrical characterization of the so-called Maxwell set, wave fronts and caustics.

As we mentioned before, we study the integrability of the constrained Hamiltonian system provided by the necessary conditions of the Maximum Principle. After analyzing a group action that leaves our system invariant, we discuss the integrability of the system by explicitly carrying out an integration procedure that applies to the known low dimensional cases. We show that the resulting Hamiltonian system can be explicitly integrated in terms of hyperelliptic functions. The integrability of left invariant Hamiltonian systems of the class we study here, was suggested by R. Montgomery et al in [9] in connection with the integrability question of the geodesic flow on Carnot groups.

Apart from this introduction, this paper contains 5 sections. In Section 2 we discuss the Gaveau-Brockett-Sussmann problem. We study certain group actions that leave the left invariant control system invariant. The results in this section extend the actions of semidirect products of tori and base manifolds that we use for the higher dimensional Heisenberg in [16]. In section 3 we set the optimal control problem, establish our notation and argue about the general sub-Riemannian geodesic problem in nilpotent Lie groups of order one. In section 4, we set the symplectic structure underlying the problem and apply the Maximum Principle to discuss the general properties of extremal curves. In section 5 we carry out an explicit integration process for the normal extremals, exhibiting a complete set of integral of motion in involution. Finally in section 6, we discuss, within this general setting, some aspects of the low dimensional cases known in the literature.

2 Invariance of the Gaveau-Brockett-Sussmann system

In this section we study a left invariant control system on the nilpotent Lie group $\mathbb{R}^n \times \mathrm{SO}_n$, such system provides a model for the left invariant optimal control problem on nilpotent Lie groups that we consider in this paper. This system was first study by B. Gaveau [1] then by R. Brockett [2] and later on by W. Liu and H. Sussmann [5].

Let us consider the control system in $\mathbb{R}^n \times \mathrm{SO}_n$ given by the following equations

$$\begin{aligned} \dot{x}_i &= u_i, \quad i = 1, 2, \ldots, n \\ \dot{z}_{ij} &= u_i x_j - u_j x_i, \qquad i, j = 1, 2, \ldots, n, \end{aligned}$$

these equations can be assembled as follows

$$(\dot{x}, \dot{z}) = \sum_{i=1}^{n} u_i \ (e_i, x^T \otimes e_i - e_i^T \otimes x). \tag{1}$$

where $\{e_i\}$ denotes the canonical basis of $\mathbb{R}^n$, $\dot{z}$ is a skew-symmetric matrix and $\otimes$ denotes the Kronecker product of vectors in $\mathbb{R}^n$.

After identifying the group SO_n with its Lie algebra $\mathfrak{so}_n$ we shall denote as H_{n+1}, the Lie group with underlying manifold $\mathbb{R}^n \times \mathfrak{so}_n$ and group law:

$$(x_1, z_1) \odot (x_2, z_2) = (x_1 + x_2, z_1 + z_2 - (x_1^T \otimes x_2 - x_2^T \otimes x_1)).$$

The group H_{n+1} has dimension $\frac{n(n+1)}{2}$, and the vector fields on the right-hand side of equation (1) turns out to be left invariant.

In fact, by considering coordinates $(x_1, \ldots, x_n, z_{12}, \ldots, z_{(n-1)n})$, it can be easily verified that a basis of left invariant vector fields is given by

$$X_i = \partial x_i + \sum_{j \neq i}^{n} x_j \partial z_{ij}, \quad i = 1, \ldots, n.$$

together with the following Lie brackets

$$[X_i, X_j] = -2\partial z_{ij} =: X_{ij}, \quad 1 \leq i < j = 2, \ldots, n.$$

Observe that the vector field ∂x_i corresponds to the element in the tangent space $T_0\mathbb{R}^n$ given by e_i, whereas ∂z_{ij} corresponds to the element in $T_0\,\mathfrak{so}_n$ given by the matrix $e_i^T \otimes e_j - e_j^T \otimes e_i$. Furthermore, all the vector fields X_{ij} are central elements.

Remark 2.1 *The group* H_3 *corresponds to the three dimensional Heisenberg group with the usual Heisenberg translation. Furthermore, the subalgebras* $\{X_i, X_j, X_{ij}\}$ *for* $i \neq j$ *are also isomorphic to the three dimensional Heisenberg algebra.*

2.1 Invariance of the system

We describe now a group action on H_n for which system (1) remains invariant. For this purpose it results more appropriate to write the system by means of 1-forms. More precisely, the system is defined by the kernel of the Pfaffian system

$$\omega(x, z) := dz - (x^T \otimes dx - dx^T \otimes x), \tag{2}$$

where $dz = (dz_{ij})$ and $dx = (dx_1, \ldots, dx_n)$

First observe that the orthogonal group SO_n acts by automorphisms on the group H_{n+1}, such action is defined by means of the mapping

$$\mathrm{SO}_n \times \mathrm{H}_{n+1} \longrightarrow \mathrm{H}_{n+1}, \quad (R, (\alpha, A)) \mapsto (\alpha, A)^R := (\alpha R, R^T A R). \tag{3}$$

Therefore we can consider the semidirect product $\mathrm{SO}_n \ltimes \mathrm{H}_{n+1}$ with elements $(R, (x, z))$ and group law defined as follows

$$(R, (\alpha, A)) * (S, (\beta, B)) := (RS, (\alpha, A)^S \odot (\beta, B)).$$

In order to simplify the notation, we shall write $a \wedge b = a \otimes b - b \otimes a$, for all $a, b \in \mathbb{R}^n$, therefore $\omega(x, z) = dz - x \wedge dx$. Furthermore, for any $S \in \mathrm{SO}_n$ and any $x, \alpha \in \mathbb{R}^n$ the following holds

$$S^T (x \wedge \alpha) S = S^T (x^T \otimes \alpha - \alpha^T \otimes x) S = (xS) \wedge (\alpha S), \tag{4}$$

we have then the following result.

Theorem 2.2 *The mapping*

$$((R, (\alpha, A)), (x, z)) \mapsto (x, z)^{(R,(\alpha,A))} := (x, z)^R \odot (\alpha, A) \tag{5}$$

defines an action of the semi-direct product $\mathrm{SO}_n \ltimes \mathrm{H}_{n+1}$ *on the group* H_{n+1}. *Furthermore, the coset of* $\omega(x, z)$ *under this action, coincides with its coset under the action of* SO_n *given by (3).*

Proof. Observe first that $(x, z)^{(I,(0,0))} = (x, z)$. We claim that

$$\left((x, z)^{(R,(\alpha,A))} \right)^{(S,(\beta,B))} = (x, z)^{(R,(\alpha,A)) * (S,(\beta,B))},$$

for all $(x, z) \in \mathrm{H}_{n+1}$ and for all $(R, (\alpha, A))$, $(S, (\beta, B)) \in \mathrm{SO}_n \ltimes \mathrm{H}_{n+1}$. In fact, equation (4) implies

$$\begin{aligned}
\left((x,z)^{(R,(\alpha,A))}\right)^{(S,(\beta,B))} &= \left((x,z)^{R}\odot(\alpha,A)\right)^{(S,(\beta,B))}\\
&= \left((xR+\alpha, R^T zR + A - (xR)\wedge\alpha)\right)^{(S,(\beta,B))}\\
&= (xRS+\alpha S+\beta, S^T R^T zRS + S^T AS\\
&\quad -S^T((xR)\wedge\alpha)S + B - (xRS)\wedge\beta - (\alpha S)\wedge\beta)\\
&= (xRS+\alpha S+\beta, S^T R^T zRS + S^T AS\\
&\quad -(xRS)\wedge(\alpha S) + B - (xRS)\wedge\beta - (\alpha S)\wedge\beta)\\
&= (x,z)^{(RS,(\alpha S+\beta, S^T AS + B - (\alpha S)\wedge\beta))}\\
&= (x,z)^{(RS,(\alpha S, S^T AS)\odot(\beta,B))}\\
&= (x,z)^{(R,(\alpha,A))*(S,(\beta,B))}.
\end{aligned}$$

as claimed. We conclude then that the action is well defined. Now let $(R,(\alpha,A))$ be an arbitrary but fixed element in $\in \mathrm{SO}_n \ltimes \mathrm{H}_{n+1}$, again equation (4) yields

$$\begin{aligned}
\omega((x,z)^{(R,(\alpha,A))}) &= \omega((x,z)^R\odot(\alpha,A))\\
&= \omega(xR+\alpha, R^T zR + A - (xR)\wedge\alpha)\\
&= d(R^T zR + A - (xR)\wedge\alpha) - (xR+\alpha)\wedge d(xR+\alpha)\\
&= R^T dzR - (xR)\wedge(dxR)\\
&= d\ (R^T zR) - (xR)\wedge(d(xR))\\
&= \omega((x,z)R)
\end{aligned}$$

as claimed.

$\square$

Remark 2.3 *This result extends the invariance, established in our previous work [16], of the Heisenberg 1-form in* $\mathbb{R}^{2n+1}$ *defined as* $\omega((\vec{x},z) = dz - \langle\vec{x}, d\vec{x}J\rangle$ *with* $(\vec{x},z)\in\mathbb{R}^{2n+1}$ *and* J *the canonical symplectic matrix.*

Remark 2.4 *The semidirect product has a corresponding analog at the level of the Lie algebra as a semidirect sum whose generators can also be constructed in a similar way.*

Remark 2.5 *The invariance of the system with respect to the action of the orthogonal group* SO_n *given by the map (3), can be explicitly written. In fact, writing the system (1) as*

$$\begin{aligned} \dot{x} &= \mathbf{u}, \\ \dot{z} &= x \wedge \mathbf{u}, \end{aligned}$$

with $\mathbf{u} = (u_1, \ldots, u_n)$*, and taking* $\xi = xR$*,* $\mathbf{v} = \mathbf{u}R$ *and* $\theta = R^T z R$*, we have that*

$$\begin{aligned} \dot{\xi} &= \mathbf{v}, \\ \dot{\theta} &= \xi \wedge \mathbf{v}. \end{aligned}$$

3 The optimal control problem

Let G be a 2-step nilpotent Lie group of dimension $n(n+1)/2$, and let $\Delta = \{X_1, \ldots, X_n\}$ be a rank-n, bracket generating and left invariant distribution on G. We consider the left invariant optimal control problem on G with dynamics

$$\dot{g}(t) = u_1(t) X_1(g(t)) + \cdots + u_n(t) X_n(g(t)), \tag{6}$$

and cost functional

$$\Lambda(g, \mathbf{u}) = \frac{1}{2} \int (u_1^2 + \cdots + u_n^2)\ dt. \tag{7}$$

The family of admissible controls is denoted as $\mathcal{U}$ and it consists of the measurable and bounded functions $t \mapsto \mathbf{u} = (u_1(t), \ldots, u_n(t))$.

We assume that $\{X_1, \ldots, X_n\}$ is a *nilpotent basis* for Δ, with order of nilpotency one, that is to say, $ad^k_{X_i}(X_j) = 0, \quad \text{for all} \quad k > 1$. Furthermore, we assume that $X_1, \ldots, X_n$ together with the non-zero Lie brackets

$$X_{ij} := [X_i, X_j], \quad 1 \le i < j = 2, \ldots, n, \tag{8}$$

determine a basis of left invariant vector fields for the Lie algebra $\mathfrak{g}$ of the group G.

As in the last section, we observe that all the vector fields $[X_i, X_j]$ are central, and also that the subalgebras $\{X_i, X_j, X_{ij}\}$ for $i \neq j$ are all isomorphic to the three-dimensional Heisenberg algebra.

3.1 The sub-Riemannian geodesic problem

The optimal control problem under consideration can be formulated as a sub-Riemannian geodesic problem on the group G. The geometry underlying the problem is of relevant importance, since nilpotent groups are natural generalization of the three dimensional Heisenberg group which is recognized as the prototype of sub-Riemannian geometry.

The sub-Riemannian structure on the group G, and the corresponding geodesic problem is naturally defined. First, by declaring the vectors $X_1(g), \ldots, X_n(g)$ orthonormal, one defines for each $g \in G$, an inner product $\langle \cdot, \cdot \rangle_{\Delta(g)}$ on the plane $\Delta(g) = \text{span}\,\{X_1(g), \ldots, X_n(g)\}$, that varies smoothly with respect to g. Then an absolutely continuous curve $g : [0, T_g] \to G$ is by definition $\Delta-$*horizontal* or *admissible*, provided $\dot{g}(t) \in \Delta(g)$, almost everywhere. The sub-Riemannian length of a $\Delta-$horizontal curve $g(t)$ is then defined as follows

$$\ell(g) = \int_0^{T_g} \|\dot{g}(t)\| \; dt.$$

The fact that the left invariant distribution Δ is bracket-generation guarantees that any two elements $g_i, g_f \in G$ can be connected by a $\Delta-$horizontal curve (Chow-Rashevskii's theorem), and consequently the sub-Riemannian distance

$$d(g_i, g_f) = \inf\{\ell(g) \mid g : [0, T_g] \to G, \text{is horizontal and } g(0) = g_i, g(T_g) = g_f\}$$

is well defined and finite.

The sub-Riemannian geodesic problem consists then in the minimization of the functional ℓ in the class of $\Delta-$horizontal curves. For $\Delta-$horizontal curves parameterized by arc-length, the problems of minimizing the functional ℓ and the functional of energy

$$\mathcal{E}(g) = \int_0^{T_g} \|\dot{g}(t)\|^2 \; dt,$$

are equivalent.

It turns out more convenient to deal with the energy minimization problem, since in that case most of the calculations are simplified. In particular, the fact that $\dot{g}(t) \in \Delta(g(t))$ almost everywhere, implies that there is a measurable and bounded function $t \mapsto (u_1(t), \ldots, u_n(t))$ (an admissible control), such that $\dot{g}(t) = u_1(t)X_1(g(t)) + \ldots + u_n(t)X_n(g(t))$ almost everywhere, and consequently the integrand in the functional $\mathcal{E}$ writes as follows

$$\begin{aligned} \|\dot{g}(t)\|^2 &= \langle u_1(t)X_1(g(t)) + \cdots + u_n(t)X_n(g(t)), \\ &\quad\; u_1(t)X_1(g(t)) + \cdots + u_n(t)X_n(g(t))\rangle \\ &= u_1^2(t) + \cdots + u_n^2(t) = 1. \end{aligned}$$

Since the arc-length parameterization for admissible curves is alway possible, we can conclude that the geodesic sub-Riemannian problem associated to the sub-Riemannian structure (G, Δ) is equivalent to the time optimal control problem associated to the system (6) under the following condition for the controls:

$$u_1^2(t) + \ldots + u_n^2(t) = 1. \tag{9}$$

4 Extremal trajectories for the adjoint system

As customary in differential geometric optimal control theory, we associated to our system another system defined at the level of the cotangent bundle, that involves linearly the control parameters, such a system is usually known as the adjoint system and its corresponding solutions are usually called extremal curves. A fundamental result in optimal control, the Pontryagin Maximum Principle, implies that optimal solutions are necessarily projections of extremal curves which further minimize in $\mathcal{U}$ the system Hamiltonian.

The adjoint system is defined by means of a control dependent Hamiltonian defined in the cotangent bundle of the group G. In order to carry out this formulation and fix our notation, we shall quote some well known results in the symplectic geometry underlying this setting, for details see for instance [26].

4.1 Lie-Poisson structure of the system

The canonical symplectic form Ω associates to each function H on the cotangent bundle T^*G its *Hamiltonian vector field* H^{Ham} by means of the formula

$$dH_\xi(v) = \Omega(v, H^{Ham}(\xi)),\ v \in T_\xi T^*G,\ \xi \in T^*G,$$

in this manner, each left invariant vector field X on the group G is canonically lifted to its Hamiltonian vector field by defining its *Hamiltonian function* $\mathcal{H}_X$ through the formula

$$\mathcal{H}_X(\xi) = \xi(X(g)),\ g \in G \quad \xi \in T_g^*G.$$

The Poisson bracket of two arbitrary Hamiltonians η_1, η_2 is given as follows

$$\{\eta_1, \eta_2\}(\xi) = \frac{d\eta_1}{dt}(\exp t\eta_2^{Ham}(\xi))\Big|_{t=0}, \quad \xi \in T^*G.$$

Lie and Poisson brackets are closely related, since if η_1 and η_2 are the Hamiltonians corresponding to the vector fields Y_1 and Y_2, then

$$\mathcal{H}_{[Y_1,Y_2]} = \{\eta_1, \eta_2\}. \tag{10}$$

The left translation on $g \mapsto \lambda_g$ on the group G, determines the trivializations

$$G \times \mathfrak{g} \simeq TG, \quad G \times \mathfrak{g}^* \simeq T^*G,$$

of the tangent and cotangent bundles by means of the mappings $(g, X) \mapsto d_e\lambda_g X$ and $(g, p) \mapsto d_e\lambda_{g^{-1}}(g)^*p$, respectively. Therefore the following representation for the double bundle TT^*G is obtained

$$TT^*G \simeq (G \times \mathfrak{g}^*) \times (\mathfrak{g} \times \mathfrak{g}^*).$$

In this manner, any tangent vector in TT^*G can be considered as a pair $((g,p),(X,Y^*))$ and the symplectic form Ω adopts the following form

$$\Omega_{(g,p)}((X_1,Y_1^*),(X_2,Y_2^*)) = Y_1^*(X_2) - Y_2^*(X_1) - p[X_1,X_2].$$

For any Hamiltonian function $H : T^*G \to \mathbb{R}$ Hamilton's equations are written as

$$\begin{aligned} \frac{dg}{dt} &= (d\lambda_g)(dH) && (11)\\ \frac{dp}{dt} &= -(ad^* dH)(p). && (12) \end{aligned}$$

For each $i = 1,\ldots,n$, we shall denote by H_i the Hamiltonian function corresponding to the vector field X_i, and for H_{ij} the one corresponding to the vector field X_{ij}, that is

$$\begin{aligned} H_i &= \mathcal{H}_{X_i},\ i = 1,\ldots,n \\ H_{ij} &= \mathcal{H}_{X_{ij}}. \end{aligned}$$

The commuting relations (8) together with the property (10) yield the following non-zero Poisson brackets

$$\{H_i, H_j\} = H_{ij}, \quad \text{for} \quad 1 \le i < j = 1,2,\ldots,n. \tag{13}$$

As before, all the H_{ij} with $i < j$ are central elements of the Lie-Poisson algebra $T^*\mathfrak{g}$.

The left invariance of the vector fields implies that all these Hamiltonians depend only on the second variable p. Furthermore, if $\{X_i^*, X_{ij}^*\}$ is the dual basis of the basis $\{X_i, X_{ij}\}$, then, the dual variable p can be identified with the vector

$$\begin{aligned} p &\simeq \sum_{i=1}^{n} p(X_i)X_i^* + \sum_{i<j} p(X_{ij})X_{ij}^* \\ &= \sum_{i=1}^{n} H_i X_i^* + \sum_{i<j} H_{ij} X_{ij}^* \end{aligned}$$

which in turn, because of the commutators (13), can be represented by a pair $(h, H) \in \mathbb{R}^n \times \mathfrak{so}_n$, with

$$\begin{aligned} h &= (H_1, H_2, \ldots, H_n) \\ H &= (H_{ij}) \end{aligned}$$

4.2 The Maximum Principle

Now we write the necessary conditions for optimality. Each admissible control function $\mathbf{u} \in \mathcal{U}$, determines a Hamiltonian function as follows:

$$\mathcal{H}_{\lambda,\mathbf{u}} = -\frac{\lambda}{2}(u_1^2 + \cdots + u_n^2) + u_1 H_1 + \cdots + u_n H_n,$$

where the parameter λ can be normalized to take the value 1 or 0.

Given an arbitrary admissible control function $\mathbf{u} = (u_1, \ldots, u_n) \in \mathcal{U}$ and certain fixed initial conditions, the integral curves of the Hamiltonian system (11) and (12) corresponding to the Hamiltonian $\mathcal{H}_{\lambda,\mathbf{u}}$ are called *extremal curves.* Those which correspond to $\lambda = 1$ are called *normal* extremals, whereas the ones corresponding to $\lambda = 0$ are called *abnormal.* The following is an invariant version of the classical Pontryagin Maximum Principle, see [14].

Theorem 4.1 (Pontryagin) *A solution curve $t \mapsto (g(t), \widehat{\mathbf{u}})$ of system (6), is Λ-optimal, if it is the projection of an extremal curve (g, p) along which the inequality $\mathcal{H}_{\lambda,\widehat{\mathbf{u}}} \geq \mathcal{H}_{\lambda,\mathbf{w}}$ holds a.e., for all $\mathbf{w} \in \mathcal{U}$. Furthermore, for $\lambda = 0$, the dual variable p cannot be identically zero.*

4.3 Extremal curves

For the remaining of the paper, we assume that in a certain interval $[0, T_g]$, the trajectory $t \mapsto (g(t), \widehat{\mathbf{u}}(t))$ is the projection of an extremal curve $(g(t), p(t)) = (g(t), h(t), H(t))$.

4.3.1 Abnormal extremals.

As pointed out before without loss of generality we can impose the following condition on the admissible controls

$$u_1^2 + \cdots + u_n^2 = 1. \tag{14}$$

In the abnormal case, the Hamiltonian is given as follows

$$\mathcal{H}_{0,\widehat{\mathbf{u}}} = \widehat{u}_1 H_1 + \cdots + \widehat{u}_n H_n,$$

and the optimality condition of the maximum principle yields the following constraints

$$\begin{aligned} H_1 &= 0, \\ &\vdots \\ H_n &= 0. \end{aligned} \tag{15}$$

to be satisfied along the extremal curve.

Theorem 4.2 *If $(g(t), h(t), H(t))$ is abnormal, then $h = 0$, the matrix H is non-zero constant, and the optimal control $\widehat{\mathbf{u}}$ belongs to the kernel of H. Furthermore, if H has all its eigenvalues different from zero, then for n even, $g(t) = g(0)$ for all t, whereas for n odd $g(t)$ is solution of the equation $\dot{g} = \pm X_n$.*

Proof. The constraints (15) directly imply $h = (0, \dots, 0)$, moreover, since the H_{ij} are all central elements, then along the extremal we have

$$\dot{H}_{ij} = \{H_{ij}, \widehat{u}_1 H_1 + \cdots + \widehat{u}_n H_n\} = 0,$$

therefore H is constant skew-symmetric matrix and Maximum Principle guarantees that not all its entries are equal to zero. Now by differentiating the constraints along the extremal we obtain

$$\begin{aligned}
0 &= \dot{H}_1 = \{H_1, \widehat{u}_1 H_1 + \cdots + \widehat{u}_n H_n\} = \widehat{u}_2 H_{12} + \cdots + \widehat{u}_n H_{1n} \\
&\vdots \\
0 &= \dot{H}_i = \{H_i, \widehat{u}_1 H_1 + \cdots + \widehat{u}_n H_n\} = \widehat{u}_1 H_{i1} + \cdots + \widehat{\widehat{u}_i H_{ii}} + \cdots + \widehat{u}_n H_{in} \\
&\vdots \\
0 &= \dot{H}_n = \{H_n, \widehat{u}_1 H_1 + \cdots + \widehat{u}_n H_n\} = \widehat{u}_1 H_{n1} + \cdots + \widehat{u}_{n-1} H_{n(n-1)},
\end{aligned}$$

here $\widehat{\widehat{u}_i H_{ii}}$ means that such a term is omitted. In consequence $\widehat{\mathbf{u}}$ is in the kernel of the matrix H, as claimed.

Assume now that along the extremal, H is a $n \times n$ skew-symmetric constant matrix with all its eigenvalues $\alpha_1, \dots, \alpha_n$ different from zero. If n is even then there is a matrix $R \in \mathrm{SO}_n$ such that the matrix $D = R^T H R$ writes as the $2 \times 2-$block diagonal matrix

$$\mathrm{diag}(D_1 \ \cdots \ D_n),$$

with

$$D_i = \begin{pmatrix} 0 & \alpha_i \\ -\alpha_i & 0 \end{pmatrix}.$$

But then, H has trivial kernel and the only solution to the system $H\widehat{\mathbf{u}} = 0$ is the trivial solution. Therefore the equation (6) reduces to $\dot{g} = 0$.

similarly, if n is odd and $\alpha, \dots, \alpha_{n-1}$ are the non-zero eigenvalues, there is again a $R \in \mathrm{SO}_n$ such that $D = R^T H R$ writes as a $2 \times 2-$block diagonal matrix with an extra column of zeros, that is,

$$\mathrm{diag}(D_1 \ \cdots D_{n-1} \ 0),$$

therefore $\widehat{\mathbf{u}} = (0, \ldots, 0, \widehat{u_n})$, and condition(14) reduces the equation (6) to $\dot{g} = \pm X_n$, as claimed.

□

Remark 4.3 *According to Liu and Sussmann's definition, [5], a strictly abnormal extremal is an abnormal extremal that is not normal. It is clear from the expressions we have founded above that the abnormal extremals in our case turn out to be non strictly abnormal.*

4.3.2 Normal extremals

Now we analyze the case of normal extremals. In this case the system Hamiltonian is given as follows

$$\mathcal{H}_{1,\mathbf{u}} = -\frac{1}{2}(u_1^2 + \cdots + u_1^2) + u_1 H_1 + \cdots + u_n H_n,$$

and the optimality condition of the maximum principle implies that

$$\begin{aligned} u_1 &= H_1 \\ &\vdots \\ u_n &= H_n \end{aligned}$$

along the extremal. The system Hamiltonian then takes the form

$$\mathcal{H} = \frac{1}{2}\left(H_1^2 + \cdots + H_n^2\right),$$

and differentiating the $H_{i's}$ along the extremal we get

$$\begin{aligned} \dot{H}_1 &= \{H_1, \mathcal{H}\} = H_2 H_{12} + H_3 H_{13} + \cdots + H_n H_{1n} \\ \dot{H}_2 &= \{H_2, \mathcal{H}\} = H_1 H_{21} + H_3 H_{23} + \cdots + H_n H_{2n} \\ &\vdots \\ \dot{H}_n &= \{H_n, \mathcal{H}\} = H_1 H_{n1} + H_2 H_{n2} + \cdots + H_{n-1} H_{(n-1)n}, \end{aligned}$$

or equivalently $\dot{h} = H\ h$. As before, since the H_{ij} are central elements, the skew-symmetric matrix H remains constant along the extremal. In fact

$$\begin{aligned} \dot{H}_{ij} &= \{H_{ij}, \mathcal{H}\} = \frac{1}{2}\{H_{ij}, H_1^2 + \cdots + H_n^2\} \\ &= H_1\{H_{ij}, H_1\} + \cdots + H_n\{H_{ij}, H_n\} = 0. \end{aligned}$$

In summary, we can write the following result for the normal extremals

Theorem 4.4 *If $(g(t), h(t), H(t))$ is a normal extremal, then*

$$\frac{dg}{dt} = H_1 X_1(g) + \cdots + H_n X_n(g), \quad (16)$$

$$\frac{dp}{dt} = (\dot{h}, \dot{H}) = (Hh, \mathbf{0}), \quad (17)$$

5 Integration of the adjoint system

We analyze now the integrability of the adjoint system, we concentrate in the normal case since, as pointed out above, the abnormal extremals are not strictly abnormal. Integration of equation (17) shall provide the complete solution of the problem, since the optimal controls are explicitly given: $\widehat{\mathbf{u}} = h$

Lemma 5.1 *The system Hamiltonian $\mathcal{H}$ is constant along the extremal.*

Proof. In fact, a direct differentiation along the extremal yields

$$\dot{\mathcal{H}} = H_1 \{H_1, \mathcal{H}\} + \cdots + H_n \{H_n, \mathcal{H}\} = h^T H h = 0$$

since H is skew-symmetric. □

The n initial conditions $h(0) = (H_1(0), \ldots, H_n(0))$, together with the $n(n-1)/2$ constants $H_{k_{ij}}$ provide a complete set of integrals of motion for the system (17) that guarantee the integrability of the system by quadratures, furthermore we can explicitly calculate the solutions.

Theorem 5.2 *The adjoint system (17) is integrable by quadratures and its integral curves are given by*

$$h(t) = e^{tH} h_0.$$

Since H is skew-symmetric its eigenvalues are purely imaginary and they appear in $\pm$ pairs or are zero. The exponential e^{tH} then can be explicitly calculated applying the classical Sylvester's theorem of linear algebra.

Let $f : \mathbb{C} \to \mathbb{C}$ be an arbitrary analytic function and let A be an arbitrary $n \times n$ square matrix, having all its eigenvalues in the open disk $|z| < \rho$, then the matrix

$$f(A) = \frac{1}{2\pi i} \int_{|z|=\rho} f(z)(zI - A)^{-1} \, dz$$

is well defined and its entries are given as follows

$$f_{ij} = \frac{1}{2\pi i} \int_{|z|=\rho} f(z) \langle e_i^T, (zI - A)^{-1} e_j \rangle \, dz.$$

For each i, j the function $z \mapsto f(z)\langle e_i^T, (zI-A)^{-1}e_j\rangle$ is meromorphic in the disk $|z| < \rho$, and its poles are precisely the eigenvalues of A. Furthermore, if for a non-singular matrix R we have that

$$RAR^{-1} = \operatorname{diag}(A_1 \cdots A_p)$$

with the $A_{i'}s$ square blocks of appropriate size, then

$$f(A) = R^{-1}\operatorname{diag}(f(A_1)\cdots f(A_p))R.$$

Of course, the block A_p might be the zero matrix, but then $f(A_p) = I_p$.

Let $\sigma(A) = \{\lambda_1, \dots, \lambda_p\}$ be the set of eigenvalues of A, and for each k, let ν_k denote the multiplicity of the eigenvalue λ_k, the characteristic polynomial of A is then written as $p(z) = \Pi(z-\lambda_i)^{\nu_i}$, furthermore,

$$f(A) = \frac{1}{2\pi i}\int_{|z|=\rho}\frac{f(z)}{p(z)}\operatorname{Adj}(zI-A)\;dz$$

by applying residue theorem at each entry we obtain sylvester's formula

$$f(A) = -\sum_{i=1}^{p}\frac{1}{(\nu_i-1)!}\cdot\frac{d^{\nu_i-1}}{dz^{\nu_i-1}}\left((z-\lambda_i)^{\nu_i}\frac{f(z)}{p(z)}\operatorname{Adj}(zI-A)\right). \tag{18}$$

In particular if $\nu_i = 1$ for all i, that is, if all the eigenvalues of A are different, then equation (18) yields

$$f(A) = \sum_{i=1}^{p} f(\lambda_i)\prod_{j\neq i}\frac{A-\lambda_j I}{\lambda_i-\lambda_j}$$

Since this formula is valid for any analytic function, we can use it for calculate the exponential e^{tH}.

6 Low dimensional cases

We now specialize the results to the cases $n = 2$ and $n = 3$

(2,3)-Case, the Heisenberg algebra. This case corresponds to $n = 2$ and has been widely studied since the pioneering paper of R.W. Brockett [2]. Within the framework we have presented here, it corresponds to the three dimensional Lie algebra given by

$$\{\eta_1, \eta_2\} = \eta_3.$$

For the normal extremals, the adjoint system writes as follows

$$\begin{aligned} \dot{H}_1 &= H_2 H_{12} \\ \dot{H}_2 &= -H_1 H_{12} \\ \dot{H}_{12} &= 0. \end{aligned} \tag{19}$$

Thus

$$H = \begin{pmatrix} 0 & H_{12} \\ -H_{12} & 0 \end{pmatrix}, \tag{20}$$

and therefore for $H_{12} \neq 0$,

$$h = \begin{pmatrix} \cos(H_{12}t) & \sin(H_{12}t) \\ -\sin(H_{12}t) & \cos(H_{12}t) \end{pmatrix} \tag{21}$$

If $H_{12} \to 0$, then the extremal degenerates into a single point.

(3,6)-Case. This case corresponds to $n = 3$, it has been studied by Brockett, Liu and Sussman, Myasnichenko and others. It consists of the six dimensional nilpotent Lie algebra given by

$$[\eta_1, \eta_2] = \eta_{12}, \;\; [\eta_3, \eta_1] = \eta_{31}, \;\; [\eta_2, \eta_3] = \eta_{23}.$$

For the normal extremals the adjoint system writes as follows

$$\begin{aligned} \dot{H}_1 &= H_{12} H_2 - H_{31} H_3 \\ \dot{H}_2 &= -H_{12} H_1 + H_{23} H_3 \\ \dot{H}_3 &= H_{31} H_1 - H_{23} H_2, \\ \dot{H}_{12} &= 0 = \dot{H}_{23} = \dot{H}_{31}. \end{aligned}$$

The matrix H is 3×3 and has eigenvalues $\{0, \pm i\lambda\}$ with

$$\lambda = \sqrt{H_{12}^2 + H_{23}^2 + H_{31}^2}.$$

Thus,

$$h = \left(1 + \sin(\lambda t)\frac{H}{\lambda} + (1 - \cos(\lambda t))\frac{H^2}{\lambda^2}\right) h_0.$$

Acknowledgments

The present paper was written while the first author (Monroy-Pérez) was on sabbatical leave at INRIA, Unité de Recherche Sophia-Antipolis, France. The author thanks Prof. J.B. Pomet and the rest of the team MIAOU for their support. He was also partly supported by the CONACYT's grant for sabbatical periods abroad.

References

[1] GAVEAU B., Principe de moindre action, propagation de la chaleur et estimées sous elliptiques sur certains groupes nilpotents, *Acta Mathematica*, **30**, pp. 94-153, (1977).

[2] BROCKETT R.W., Control theory and singular Riemannian geometry, *New directions in applied mathematics*, Hilton, P.J. and Young G.S. Eds., Springer-Verlag, (1981).

[3] MONTGOMERY R., Abnormal minimizers, *SIAM J. Control Optim.*, **32**, No. 6, pp. 1605-1620, (1994).

[4] KUPKA I., Abnormal extremals, preprint (1992).

[5] LIU W., SUSSMANN H., Shortest paths for sub-Riemannian metrics on rank-two distributions, MEMOIRS OF THE AMERICAN MATHEMATICAL SOCIETY, **118**, No.564, (1995).

[6] R. MONTGOMERY, *A tour of Subriemannian geometries, their geodesics and applications*, Mathematical surveys and monographs, Vol. 91, American Mathematical Society, (2002).

[7] HERMES H., LUNDELL A., SULLIVAN D., Nilpotent bases for distributions and control systems *J. Differential equations*, **55**, 385-400, (1984).

[8] BELLAÏCHE A., The tangent space in sub-Riemannian geometry. In *SubRiemannian geometry*, Bellaïche A., Risler J.-J., Birkhäuser, (1996).

[9] R. MONTGOMERY, M. SHAPIRO AND A. STOLIN, A non-integrable sub-Riemannian geodesic flow on a Carnot group, *J.Dynam. Control Systems*, **3**, No. 4, 519-530, (1997).

[10] CARTAN E., Lès sous-groupes des groupes continus de transformations, *Annales de l'Ecole Normale*, **25**, 1908,57-194. Reprinted in Œuvres completes Partie II, 2nd. ed, Editions du Centre National de la Recherche Scientifique, pp. 710-856, (1984).

[11] R.L. BRYANT, S.S. CHERN, R.B. GARDNER, H.L. GOLDSCHMIDT, AND P.A. GRIFFITHS, *Exterior Differential systems*, Springer Verlag, (1991).

[12] R. GARDNER, The method of equivalence and Applications, *SIAM-CBMS, Regional Conf. Ser. in Appl. Math.*, **58**, Philadelphia, (1989).

[13] R.M. Murray, Nilpotent basis for a class of non-integrable distributions with applications to trajectory generation for nonholonomic systems , *Mathematics of controls, signals and systems*, **7**, No.1, pp. 58-75, (1994).

[14] L.D. Pontryagin, V. Boltiansky, R. Gamkrelidze, and E. Mishenko, *The mathematical theory of optimal processes theory*, Wiley, New York (1962).

[15] A.M. Vershik and V. Ya. Gershkovich, Nonholonomic dynamical systems, geometry of distributions and variational problems. In: *Enyc. Mathematical Sciences,* Vol. 16; *Dynamical systems VII*, Springer-Verlag, (1991).

[16] A. Anzaldo-Meneses and F. Monroy-Pérez, Optimal control in the Heisenberg group, *J. Dynam. Control Systems*, **5**, No. 4, 473-499, (1999).

[17] B. Bonnard, M. Chyba and I. Kupka Nonintegrable geodesics in SR-Martinet geometry. Differential geometry and control (Boulder, CO, 1997), 119–134, *Proc. Sympos. Pure Math.*, **64**, Amer. Math. Soc., Providence, RI, (1999).

[18] Agrachev, A. A. ; Charlot, G.; Gauthier,; Zakalyukin, V. M., On sub-Riemannian caustics and wave fronts for contact distributions in the three-space. *J. Dynam. Control Systems*,**6**, no. 3, pp. 365-395, (2000).

[19] Charlot, G. , Quasi-contact S-R metrics: normal form in $\mathbb{R}^{2n}$, wave front and caustic in $\mathbb{R}^4$. *Acta Appl. Math.*,**74**, no. 3, pp.217-263. (2002).

[20] A.M. Vershik and O. Granichina, Reduction of nonholonomic variational problems to the isoperimetric problems and connections in principal bundles (Russian) Mat. Zametki **49**, 37-44, (1991).

[21] Y.L. Sachkov, Symmetries of Flat Rank Two Distributions and Sub-Riemannian Structures, *Rapport de Recherche No.151, Université de Bourgogne, Laboratoire de Topologie*, Dijon, France, (1998).

[22] E. Cartan, Lès systemes de Pfaff a cinque variables et lès equations aux derivees partielles du second ordre, *Ann. Sci. Ecole Normale*, **27**, No.3, 109-192, (1910).

[23] R. Brockett and L. Dai, Nonholonomic kinematics and the role of elliptic functions in constructive controllability. In: *Nonholonomic motion planning*, Z.Li and J. Canny Eds., Kluwer Academic publishers, Boston, 1-21, (1993).

[24] R. Bryant and L. Hsu, Rigidity of integral curves of rank 2 distributions. *Invent. Math.*, 114, 435-461, (1993).

[25] O. Myasnichenko, Nilpotent (3,6) sub-Riemannian problem, *J. Dynam. Control Systems*,**8**, no. 4, pp.573–597, (2002).

[26] V. Jurdjevic, *Geometric control theory*, Cambridge University Press, (1996).

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399